

ANDREW ESHLEMAN

ALTERNATIVE POSSIBILITIES AND THE FREE WILL DEFENCE

I

The free will defence attempts to show that belief in an omnibenevolent, omnipotent, and omniscient God may be rational, despite the existence of evil. At the heart of the free will defence is the claim that it may be impossible, even for an omnibenevolent, omnipotent, and omniscient God, to bring about certain goods without the accompanying inevitability, or at least overwhelming probability, of evil. The good in question is the existence of free agents, in particular, agents who are sometimes free with respect to morally significant actions and who are thereby responsible, at least in part, for those actions and the personal character which is a function of and exhibited in those actions. The free will defender contends that if an agent is to be truly responsible for her actions, then she must be free to bring about both good and evil, and God cannot be blamed if such agents choose to bring about the latter rather than the former.

A number of years ago, Antony Flew objected that God was not forced to choose between creating free agents who might act wrongly and not creating a world with free agents. Instead, God could have created free agents who were wholly good, i.e. who always acted rightly.¹ Freedom and responsibility, Flew argued, are compatible with one's actions being causally determined by God, thus it was within God's power to create agents who were both free and responsible yet causally determined to always act rightly.

In response, proponents of the free will defence criticized Flew's conditional analysis of freedom – if S had chosen to do otherwise, she would have been able to do otherwise – maintaining instead that an agent's freedom consists in her ability at the time in question to both perform the action and refrain from performing the action. Acting freely, on this libertarian view, is incompatible with one's actions being determined by God, for an agent

¹ Antony Flew, 'Divine Omnipotence and Human Freedom', in *New Essays in Philosophical Theology*, ed. Antony Flew and Alasdair MacIntyre (London: SCM Press Ltd, 1955), pp. 144–169. See also J. L. Mackie, 'Evil and Omnipotence', *Mind* 64 (1955), 200–212. Some, notably Alvin Plantinga, have thought that Flew and Mackie were offering different objections. As Plantinga sees it, Flew's objection depends on the claim that acting freely is compatible with causal determinism, while Mackie's objection does not. See Alvin Plantinga, *God, Freedom, and Evil* (Grand Rapids: Wm. B. Eerdmans Publishing Co., 1977), pp. 31–33.

whose actions are determined by God is unable to do anything other than what God determines her to do. Thus, it was not possible for God to create a world with free creatures who always act rightly, for the actions of a free agent are, in a robust sense, up to her.²

It is important to note that both the free will defender and Flew (along with most incompatibilists and compatibilists in the recent history of the free will debate) share the belief that an agent is free, only if she has access to some alternative course of action and that this sort of freedom is a necessary condition of moral responsibility. Harry Frankfurt has called this the ‘principle of alternate possibilities’: ‘A person is morally responsible for what he has done only if he could have done otherwise.’³ The crux of the disagreement between the free will defender and Flew is whether ‘could have done otherwise’ should be interpreted in a manner consistent with the truth of determinism. However, Frankfurt has introduced a purported counterexample to the principle of alternate possibilities which, if successful, undercuts this debate.⁴ ‘Frankfurt-style examples’⁵ appear to show that one can act freely and thus be morally responsible for an action, even though it is not the case that one could have done otherwise.⁶

If Frankfurt-style examples are counterexamples to the principle of alternate possibilities, then it seems that Flew’s objection can be resurrected, *without relying on his controversial conditional analysis of freedom*. That is, if it can be shown that freedom to do otherwise is not required for free action or responsibility, this would be a strong reason for thinking that some version of compatibilism (one which no longer required a conditional analysis of the ability to do otherwise) was true.⁷ And if freedom and responsibility are compatible with determinism, then God could have created a world inhabited by free and responsible agents who nevertheless were causally determined to always act rightly.

² Plantinga has no doubt provided the most rigorous development of this line of response. According to Plantinga, if what a free agent does is truly up to him, then which possible worlds God is able to actualize is, at least in part, a function of what free agents choose to do. See *God, Freedom, and Evil* and *The Nature of Necessity* (New York: Oxford University Press, 1974), ch. 9.

³ Harry G. Frankfurt, ‘Alternate Possibilities and Moral Responsibility’, *The Journal of Philosophy* 66 (December 1969), 829. ⁴ *Ibid.* pp. 832–837. Below I provide a version of such an example.

⁵ Since its introduction, Frankfurt’s example has served as a model for an array of similar examples, thus it proves useful to have a way of referring to, not only Frankfurt’s original example, but all those that are relevantly similar.

⁶ If one is convinced that acting freely cannot be understood apart from its association with alternatives, then what Frankfurt-style examples show, if they are successful, is that acting freely is not a necessary condition of moral responsibility. On the other hand, one might say that the examples, if successful, show that acting freely need not be understood in terms of access to alternatives and thus maintain that free action is a necessary component of moral responsibility. I have chosen this latter approach.

⁷ This is simply because the main obstacle to the acceptance of compatibilism has traditionally been that agents whose actions were determined would not have access to the kind of alternatives necessary for freedom and responsibility. See my section III for a version of incompatibilism which construes the threat of determinism in a different manner.

Frankfurt-style examples have been widely discussed in the literature on free will and moral responsibility, yet their significance for the free will defence is rarely acknowledged.⁸ This paper seeks to remedy, at least in part, this neglect. In order to evaluate whether Frankfurt-style examples provide a basis for an objection, like that of Flew's, to the free will defence, I first introduce a Frankfurt-style example for the sake of discussion, then I address two strategies incompatibilists have employed in order to show that such examples fail to demonstrate that alternative possibilities are not required for freedom and moral responsibility. I will argue that these strategies are inadequate. Thus I claim there is reason to think that Frankfurt-style examples provide grounds for a Flew-like objection to the free will defence.

Furthermore, it seems that Frankfurt-style examples may pose a threat to the free will defence, even if one's assessment of the incompatibilist's responses to such examples is more favorable than my own. In the final section of the paper I briefly pursue this suggestion.

II

Frankfurt-style examples exhibit the following common structure: An agent decides upon a course of action and proceeds to act in a normal responsibility-conferring manner. Unbeknownst to the agent, circumstances are such that if the agent had shown signs of choosing a course of action other than the one she did, some coercive factor would come into play and guarantee that she perform her originally intended action. Significantly however, this coercive factor never comes into play.

Consider, for example, Ubiquitous Jones. Motivated by the desire to keep up with her neighbours, Jones decides and proceeds to embezzle money from her employer in order to buy a new car. Unbeknownst to Jones, her deliberations have been closely monitored by the mad doctor Green who has implanted a device in Jones's brain for monitoring her brain states and manipulating them if he so desires. Green wishes to study methods of embezzling, so had Jones shown any sign of deciding not to embezzle, Green would have intervened decisively by manipulating her neural processes, guaranteeing that she would in fact embezzle from her employer. Since Jones decides for herself to embezzle, Green does not intervene.

⁸ Frankfurt-style examples occasionally receive brief mention in discussions of the free will defence, yet authors generally fail to note the extent of the threat such examples pose if they are successful. For example, John Bishop develops an objection to the free will defence which depends indirectly on the success of the Frankfurt-style examples (Bishop seems to assume the examples are successful – he mentions them only in passing.), but he and many others seem not to notice the more direct threat such examples pose if they are indeed successful. See John Bishop, 'Compatibilism and the Free Will Defence', *Australasian Journal of Philosophy* 71 (1993), 108. The main goal of this paper is to highlight the direct threat such examples pose to the free will defence, but I conclude with an indirect argument which resembles in some respects the one offered by Bishop.

Jones appears to act freely and to be morally responsible for her act of embezzlement, yet it is clear that Green makes it impossible for her to have done otherwise. Examples like this seem to show that an agent can act freely and be morally responsible for her actions, even though she lacks access to alternatives, or freedom to do otherwise. In other words, the principle of alternate possibilities appears to be false. Moreover, since the free will defence relies on the principle of alternate possibilities, its success appears to be threatened by the Frankfurt-style examples. Again, this is so because if freedom and responsibility do not require access to alternatives, then a major obstacle to the acceptance of compatibilism will have been removed. And if freedom and responsibility are compatible with determinism, then the free will defence is vulnerable to an important objection – that God, by establishing the necessary initial conditions and natural laws, could have created a deterministic world with free agents who were responsible for their actions yet who never acted wrongly.

III

As I noted earlier, little attention has been devoted to the significance of Frankfurt-style examples for the free will defence. However, such examples have provoked numerous responses by those who wish to defend the incompatibility of moral responsibility and causal determinism. Responses to Frankfurt-style examples may be viewed as attempts to answer one of the following two questions: (1) do such examples show what they appear to show?; or (2) if they do, is this sufficient reason for thinking that moral responsibility is compatible with causal determinism? Incompatibilists have sought to answer ‘no’ to one or the other of these two questions, and the free will defender, insofar as she is an incompatibilist, will presumably wish to do the same. In other words, she will likely adopt one of the two following strategies: Strategy 1 – claim that such examples fail to demonstrate that one can be responsible for an action when one’s access to alternatives is blocked; or Strategy 2 – claim that even if the examples do show that access to alternatives is not required for freedom and moral responsibility, this does not provide adequate grounds for thinking that freedom and moral responsibility are compatible with determinism. I will discuss two versions of the first strategy in this section. The second strategy will be taken up in the following section.⁹

⁹ The responses to Frankfurt-style examples that I discuss were selected as representatives of the more general strategies just described. Although my comments and criticisms will often depend upon features of the particular responses I have chosen to discuss, I will also attempt to point out how the problems I identify in such responses may exhibit more general difficulties facing anyone who pursues a similar strategy. For a recent response to Frankfurt-style examples which differs in kind from those I discuss see James W. Lamb, ‘Evaluative Compatibilism and the Principle of Alternate Possibilities’, *Journal of Philosophy* 10 (1993), 517–27. Lamb’s objection has been successfully rebutted, I think, in John M. Fischer and Paul Hoffman, ‘Alternative Possibilities: A Reply to Lamb’, *Journal of Philosophy* 6 (1994), 321–26.

(a) *Strategy 1. Naylor's version*

A number of attempts have been made to show that Frankfurt-style examples are not genuine counterexamples to the principle of alternate possibilities. An initially promising version of this strategy is motivated by the observation that even in the most elaborate examples, it appears possible to construe the agent as having access to some sort of alternative.¹⁰ The incompatibilist strategy developed from this observation involves three steps: (a) distinguishing more carefully what kinds of things an agent may be responsible for; (b) pointing out that, in a Frankfurt-style example, there is at least one thing such that the agent is responsible for it, and it is something to which the agent had an alternative (the purpose of this step, in conjunction with the next, is presumably to help explain why the agent in such examples appears to be responsible for what she did); and (c) showing that the agent in a Frankfurt-style example is not responsible for that to which she had no alternatives (this step is crucial to dispel the initial impression that the agent is responsible, despite her lack of alternatives).

Margery Bedford Naylor has offered a version of this strategy.¹¹ First, Naylor points out that one might distinguish between being responsible for some action, A, and being responsible for doing A *on one's own*.¹² For example, a child may tell the truth about breaking the neighbour's window because she wants to tell the truth, even though she would have been compelled by her parents to tell the truth had she chosen not to do so. Naylor suggests that while it is clear in such a case that the child deserves moral credit for telling the truth *on her own*, it is perhaps not as obvious that she deserves moral credit for telling the truth, since the latter is something she would have done even if she had initially decided against it. The second step in Naylor's argument is to suggest that likewise in a Frankfurt-style case, such as the one I outlined earlier, it is clear that Jones is responsible for embezzling *on her own* but perhaps not so clear that she is responsible for the act of embezzling itself. As with the child, it is up to Jones whether to do A *on her own*, but it is not up to her whether or not to do A.¹³

Proceeding to the third step of her strategy, Naylor attempts to show that the agent in a Frankfurt-style example is not responsible for A by arguing

¹⁰ For example, as the case above suggests, Green will not intervene until Jones *shows some sign of choosing not to embezzle*. Thus it appears that Jones is capable of at least trying to choose not to embezzle. In the text I do not specifically address the sort of response which might claim that Jones's responsibility rests on this ability to try to choose, simply because its initial effectiveness seems to evaporate quickly. That is, although one might be able to make sense of the notion of trying to choose, the Frankfurt-style example can be altered so that the intervention takes place when Jones shows some sign of trying to choose. Of course, this would not eliminate every possible alternative, for one might then claim that the agent is still able to try to try to choose. In principle, this dialectic could continue indefinitely, but even at this point it is no longer clear what sort of ability one is talking about, nor what it would mean to intervene prior to its exercise.

¹¹ Margery Bedford Naylor, 'Frankfurt on the Principle of Alternate Possibilities', *Philosophical Studies* 46 (1984), 249–258.

¹² *Ibid.* p. 251.

¹³ *Ibid.* p. 252.

that an agent is not responsible for an action if it is not up to her whether or not to perform it. She begins by claiming that for any agent S, for whom it is true that it is up to him whether or not he does A *on his own*, one of two conditions might obtain:

- (1) Its being within his power not to do A, *or*
- (2) Its not being within his power not to be compelled to do A.¹⁴

According to Naylor, these conditions are both mutually exclusive and jointly exhaustive. Now in Frankfurt-style examples, it is up to the agent whether to perform A *on her own*, but it is not the case that she can refrain from performing A. Since condition (1) is thus not satisfied, condition (2) must obtain – it is not within the agent’s power not to be compelled to do A. For Naylor, this entails that the agent is compelled to do A.¹⁵ But a necessary condition of moral responsibility for an action is that the agent not be compelled to perform it. Hence, the agent in a Frankfurt-style example is not responsible for performing A.¹⁶

In sum, one might readily concede that the agent in a Frankfurt-style example is responsible for performing A *on her own*, but insist that this is in virtue of the agent having access to an alternative, i.e. not performing A *on her own* (being manipulated instead). In addition, one might argue that, contrary to first impressions, the agent in such cases is not responsible for performing A. This is something she was compelled to do, and compulsion undermines moral responsibility. If the argument stands, then Frankfurt-style examples are not cases in which it is true of any one thing both that the agent could not have done otherwise and that the agent is responsible for it. In other words, they are not genuine counterexamples to the principle of alternate possibilities.

(b) *Evaluating Naylor’s Version of Strategy 1*

The success or failure of Naylor’s argument hinges on how the notion of ‘compulsion’ is understood in step three of her strategy. Recall that the purpose of step three of the present strategy is to counter one’s initial intuition that the agent in a Frankfurt-style example is responsible for performing A, even though she could not have done otherwise. Naylor argues that the agent is not responsible for performing A because she is compelled to perform A, and she makes it clear that compulsion is to be understood as force, or coercion.¹⁷ This conclusion should strike the reader as suspicious, for while it is true in Frankfurt-style examples that the agent is not able to refrain from

¹⁴ *Ibid.* p. 253.

¹⁵ The meaning of condition (2) is less than transparent, and problems arise for Naylor on whatever reading is adopted. However, it is clear from Naylor’s use of condition (2) on p. 257 that she believes it to entail ‘S is compelled to do A’ in Frankfurt-style examples. I argue in the next section that Naylor erroneously moves from ‘It is not within S’s power not to do A’ (p. 253) to ‘S is compelled, in the sense of being coerced, to do A.’

¹⁶ *Ibid.* pp. 256–257.

¹⁷ *Ibid.* p. 252–253.

performing A, the actual sequence of events leading to her decision and subsequent action is, by stipulation, not coercive. She retains control in this sequence of events, for the coercive element (Green's intervention) never comes into play. Its role is relegated to a counter-factual, or alternative sequence.

The flaw in Naylor's reasoning becomes clear once one observes that there are at least two ways in which it might be true that an agent is compelled to do A.¹⁸ First, factors operating in the actual sequence of events leading to A might disrupt an agent's causal control of her action and thus compel the agent to perform A. For example, if Jones had shown signs of deciding otherwise, Green would have intervened as part of the actual sequence of events leading to A thereby compelling Jones to do A. Call this 'actual sequence compulsion'. Actual sequence compulsion is clearly coercive. On the other hand, the performance of an action may be made inevitable by factors which play no role in the agent's decision or subsequent action but which nevertheless make A inevitable. In other words, these factors deny the agent access to any alternative sequence but not in virtue of interfering with the causal path leading to the action. Call this 'alternative sequence compulsion'.

The difficulty with Naylor's argument is that it fails to distinguish these two forms of compulsion. Both Naylor and Frankfurt can agree that Jones is compelled to perform A, in the sense that Jones experiences alternative sequence compulsion. Unbeknownst to her, circumstances are such that it is impossible for her to refrain from embezzling from her employer. However, Naylor cannot conclude from this fact (i.e. it is inevitable that she embezzle from her employer) that Jones is not responsible for embezzling from her employer, for this is the very issue that Frankfurt-style examples are meant to decide. She tries to avoid begging the question in this obvious manner by appealing to the principle that a necessary condition of being responsible for an action is that one not be compelled, i.e. coerced, to perform it. But she then concludes that the agent in a Frankfurt-style example is coerced to perform A from the fact that she experiences alternative sequence compulsion.¹⁹ This is not acceptable for two reasons. First, Naylor's compatibilist interlocutor will surely insist that coercion should be understood in terms of actual sequence compulsion and consequently will charge that Naylor, in claiming an agent is coerced if she lacks access to alternatives, is once again assuming something Frankfurt-style examples are themselves meant to de-

¹⁸ The distinction between these two forms of compulsion is drawn from John M. Fischer's, 'Responsibility and Control', *Journal of Philosophy* 89 (January 1982), 24–40.

¹⁹ Naylor, pp. 256–257. Naylor seems to think that the agent's being coerced follows from the fact that in a Frankfurt-style example the agent lacks the power not to do A [her condition (1) fails to obtain] and it is not within the agent's power not to be compelled to do A [her condition (2) obtains]. But this is false. What does follow is that the agent could not avoid performing A and *if he had been subjected to coercion*, he would have succumbed.

cide. Secondly, to conclude that Jones is not responsible for performing A in virtue of her being coerced is surely counter-intuitive given the structure of Frankfurt-style examples. Again, by stipulation, the agent in a Frankfurt-style example arrives at her decision and performs her action in a normal responsibility-conferring manner – Green, the counter-factual intervener, plays no role in her decision and action. It is implausible to suppose that Jones’s act of embezzling funds is coerced in virtue of the mere presence of Green.

In short, Naylor’s argument trades on two notions of compulsion, and once these forms of compulsion are distinguished, her argument fails. Thus, it still appears both that it is inevitable that Jones embezzles funds from her employer and that she is responsible for doing so.

(c) *van Inwagen: Responsibility for Consequences*

One of the most powerful attempts to defend the claim that moral responsibility requires access to alternative possibilities has been made by Peter van Inwagen. Van Inwagen has offered a set of principles which he claims both are immune to Frankfurt-style counterexamples and are either versions of the principle of alternate possibilities, or close relatives.²⁰

Again, van Inwagen employs a version of Strategy 1. As a first step, he notes that our concern is often with whether an agent is responsible for the consequences of her actions. His aim is then to show that the agent’s responsibility for the consequences of her actions in a Frankfurt-style example requires that she be capable of preventing those consequences and that she is not responsible for those consequences which are inevitable.

Whether the consequences of actions should be construed as particulars or universals is a controversial matter. In order to side-step this controversy, Van Inwagen formulates one principle in which events are construed as particulars:

(PPP₁) A person is morally responsible for a certain event (particular) only if he could have prevented it.²¹

And one in which they are construed as universals:

(PPP₂) A person is morally responsible for a certain state of affairs only if (that state of affairs obtains) and he could have prevented it from obtaining.²²

Van Inwagen utilizes PPP₁ in order to show that, insofar as the agent in

²⁰ Peter van Inwagen, ‘Ability and Responsibility’, *Philosophical Review* 87 (April 1978), 201–224. I have chosen to discuss only two of the three principles that van Inwagen offers. The third, which concerns responsibility for omissions, appears to be vulnerable to counterexamples. See Robert Heintzman, ‘Incompatibilism Without the Principle of Alternative Possibilities’, *Australasian Journal of Philosophy* 64 (1986), 269–270; and Edward Wierenga, *The Nature of God* (Ithaca: Cornell University Press, 1987), pp. 77–78.

²¹ Van Inwagen, p. 206.

²² *Ibid.* p. 210.

a Frankfurt-style example is responsible for something, it is something she could have prevented. Yet before one can proceed with this second step of Strategy 1, one must first be able to ascertain whether some possible event particular would be the same as a certain actual event. That is, one needs a principle of event individuation, for according to PPP1, an agent is responsible for a given event particular only if he could have prevented it. And in order to know whether the agent could have prevented the actual event particular, one must be able to ascertain whether that same event particular would have occurred under different circumstances. Van Inwagen adopts the following criterion for individuating events:

x and y are the same event if and only if x and y have the same causes.²³

Armed with van Inwagen's criterion for individuating events, consider again Jones's responsibility for the consequence of her action (her employer's funds being embezzled), construed as a particular. Jones is responsible on this view for her employer's funds being embezzled, because she could have prevented it. That is, if she had shown signs of choosing otherwise, Green would have intervened and a different event particular would have occurred – the event particular in which her employer's funds being embezzled is due to the manipulation of Jones by Green rather than a result of Jones acting on her own. The two event particulars (characterized as (1) Jones's employer's funds being embezzled [as a result of Jones acting on her own] and (2) Jones's employer's funds being embezzled [as a result of the intervention of Green]) represent genuine alternative consequences, because they have different causes. So, if the consequences of an agent's action are construed as particulars, one can acknowledge that there is something for which the agent in a Frankfurt-style example is responsible (the event particular); however, this is not something which was inevitable for the agent. She could have prevented it from occurring.²⁴

Once again, the third step of Strategy 1 is to show that the agent in a Frankfurt-style example is not responsible for that which is inevitable. So now consider Jones's moral responsibility for the consequence of her action, construed as a universal, i.e., the state of affairs

(O) Jones's employer's funds are embezzled.

That Jones's employer's funds are embezzled is inevitable. If Jones does not embezzle the funds as a result of a choice made through normal channels of practical reasoning, the choice will be brought about via manipulation by Green. Is Jones morally responsible for this inevitable state of affairs?

²³ *Ibid.* p. 208. Van Inwagen's criterion is a modification of a principle suggested by Donald Davidson. Van Inwagen's intends the modified principle to apply not only to actual events but to merely possible events as well. For Davidson's criterion see 'The Individuation of Events', in *Essays in Honour of Carl Hempel*, ed. N. Rescher (Dordrecht, 1969), p. 225.

²⁴ Van Inwagen, p. 210.

Consider first whether Jones is responsible for the state of affairs

(L) Either Jones's employer's funds are embezzled by Jones as a result of manipulation by Green or Jones's employer's funds are embezzled by Jones acting on her own.

Van Inwagen argues that Jones is not responsible for its being the case that L.²⁵ For if Jones was responsible for such a state of affairs, then we would also have to say that Jones was responsible for its being the case that:

(M) Either Jones's employer's funds are embezzled by Jones acting on her own or $2 + 2 = 4$.

According to van Inwagen, there is no relevant difference – with respect to Jones's responsibility – between L and M, and Jones is obviously not responsible for its being the case that M. Therefore she is not responsible for its being the case that L.²⁶

If Jones is not responsible for its being the case that L, then she is also not responsible for its being the case that

(N) Either Jones's employer's funds are embezzled by Jones as a result of manipulation by Green or Jones's employer's funds are embezzled by Jones acting on her own or Jones's employer's funds are embezzled by someone else.²⁷

And since, characteristically, the same state of affairs can obtain in a number of different ways, N is identical to the state of affairs

(O) Jones's employer's funds are embezzled.²⁸

Finally, given the identity of these two states of affairs (N and O) and the fact that Jones is not responsible for its being the case that N, Jones is not responsible for its being the case that O. Hence, although it is inevitable that Jones's employer's funds are embezzled, this is not a state of affairs for which Jones is morally responsible.²⁹

So again, while there is something for which Jones is responsible and something which is inevitable, they are not the same thing. Hence, van Inwagen claims, Frankfurt-style examples fail to show that one can be morally responsible when one lacks access to alternatives.

²⁵ Of course, since the Frankfurt-style example used throughout this paper is my own, what follows is not, strictly speaking, van Inwagen's argument but rather a parallel version of the argument he offers on pp. 213–215.

²⁶ *Ibid.* pp. 213–214. The corollary state of affairs to L in van Inwagen's article is D.

²⁷ *Ibid.* p. 214.

²⁸ *Ibid.* For van Inwagen's discussion of how to determine whether two states of affairs are identical see pp. 211–212.

²⁹ Van Inwagen might grant that there is a state of affairs that Jones is responsible for, namely, that her employer's funds are embezzled by Jones acting on her own. But again this state of affairs is not inevitable for Jones.

(d) Evaluating van Inwagen's Version of Strategy 1

In my evaluation of van Inwagen's version of Strategy 1, I first make some remarks about the second step of that strategy and then argue that a crucial premise of his argument at the third stage is false.³⁰

Van Inwagen's goal at the second stage of his strategy is, like that of Naylor, to explain the widespread intuition that the agent in a Frankfurt-style example is responsible by identifying something in such cases both that the agent clearly is responsible for and to which she had an alternative. For van Inwagen this is a consequence of the agent's action, construed as an event particular, namely in the case at hand – Jones's employer's funds being embezzled (as a result of Jones acting on her own). Jones is certainly responsible for this, and it is true that there is an alternative to Jones bringing about this event particular, namely – Jones's employer's funds being embezzled (as a result of the intervention of Green). The critical question, however, is whether it is plausible to suppose that our intuitions concerning Jones's responsibility are grounded in her having access to such alternatives, i.e. the sort of alternatives required by PPP₁.³¹

John Martin Fischer has raised this question after first noting that incompatibilists traditionally have maintained a robust account of 'could have done otherwise'. That is, incompatibilists typically have insisted that responsibility requires that an agent have access to some alternative sequence where he chooses and acts otherwise in accordance with his character or on the basis of a normal process of practical reasoning.³² For example, in our scenario it normally would be necessary that Jones be able to refrain from embezzling by deciding upon and performing some action other than that of embezzling.

No such robust alternative possibility is required on van Inwagen's account. According to van Inwagen's proposal, the agent in Frankfurt-style examples is responsible for the event particular resulting from her action in virtue of her access to an alternative, even though her action in the alternative sequence does not follow from an intention she has formed on her own. Similarly, on Naylor's account the alternative which grounds the agent's responsibility is that she might have performed the action *but not on her own*. The alternative sequence in which Jones embezzles her employer's funds due to manipulation by Green is not a sequence in which Jones decides to do something different and does so. In fact, it is difficult to describe what takes place in that sequence as something which Jones does; rather it seems best

³⁰ Van Inwagen's criterion of event individuation is also controversial. For criticisms of that criterion see Fischer, 'Responsibility and Control', pp. 29–31; and Wierenga, pp. 78–81.

³¹ The concern raised here would also be relevant for PPP₂ if the universals in question were specified in sufficient detail.

³² 'Responsibility and Control', pp. 31–32. For a fuller development of this criticism, see John M. Fischer, *The Metaphysics of Free Will* (Cambridge, MA: Blackwell Publishers, 1994), pp. 140–147.

described as something which happens to her. Fischer concludes, I think rightly, that it seems implausible to suppose that our ascriptions of responsibility in Frankfurt-style examples could rest on such an attenuated notion of alternative possibilities.³³ In other words, whatever it is which grounds our intuitions about responsibility in such cases, it does not appear to be that the agent (despite appearances) in fact does have access to alternatives.

I turn now to the third stage of van Inwagen's strategy, wherein he seeks via a defence of PPP₂ to overcome the initial appearance that Jones is responsible for the state of affairs that her employer's funds are embezzled.³⁴ His argument hinges on these two claims: (1) there is no relevant difference (with respect to Jones's responsibility) between the states of affairs L and M;

(L) Either Jones's employer's funds are embezzled by Jones as a result of manipulation by Green or Jones's employer's funds are embezzled by Jones acting on her own.

(M) Either Jones's employer's funds are embezzled by Jones acting on her own or $2 + 2 = 4$.

and (2) Jones is obviously not responsible for M. There may be a relevant difference between L and M,³⁵ however I want to focus on the claim that it is obvious that Jones is not responsible for its being the case that M. Not only is this claim not obvious, as it might first appear, but I think it can be shown to be false. I suspect that van Inwagen's claim that Jones is not responsible for its being the case that M stems from his deeply held conviction that one cannot be responsible for that which is inevitable but that this is concealed by the strangeness of M. If so, like Naylor, he is implicitly relying on an assumption which itself cannot be taken for granted in Frankfurt-style examples.

To begin, note that van Inwagen's claim that Jones is not responsible for

³³ *Ibid.* Fischer's criticism seems to apply not only to van Inwagen's and Naylor's version of Strategy 1, but also to two similar versions not addressed here. See William Rowe, 'Two Concepts of Freedom', *Proceedings and Addresses of the American Philosophical Association* 61 (1987), 43–64; and Wierenga, pp. 82–85.

³⁴ Van Inwagen considers why one might think that Jones should bear responsibility for this inevitable state of affairs. One might think so, he admits, because Jones brought it about intentionally and her action was sufficient for it obtaining. But van Inwagen argues that this suggestion is problematic, for it entails that Jones is responsible for other states of affairs for which she obviously is not responsible. For example, it appears on this view that Jones would be responsible for the state of affairs that her employer is capable of being swindled, because Jones both brought about this state of affairs intentionally and her action was sufficient for its obtaining ('Ability and Responsibility', pp. 212–213). However, perhaps the point can be made another way. Revising the account slightly, one might say that the moral significance of Jones's action is not merely that it is performed intentionally and is sufficient for the obtaining of the state of affairs that her employer's funds are embezzled, but also that it partially explains why it is the case that her employer's funds are embezzled. (See Heinaman, p. 271. For a critical discussion of this view see John Martin Fischer and Mark Ravizza, 'The Inevitable', *Australasian Journal of Philosophy* 70 (1992), 388–404.)

³⁵ So argues Heinaman, pp. 272–273; William L. Rowe, 'Causing and Being Responsible for What is Inevitable', *American Philosophical Quarterly* 26 (April 1989), 158–159; and Fischer and Ravizza, 'The Inevitable', p. 401.

its being the case that *M* no doubt rests on something like the following general assumption:

- (i) If one is not responsible for its being the case that *R*, then one is also not responsible for its being the case that *P* or *R*.³⁶

I will attempt to show that assumption (i) is false in those instances in which one is responsible for its being the case that *P*.

Consider a case involving the over-determination of a state of affairs.³⁷ Kersey and Nells are members of a firing squad assigned to execute a prisoner. Both take their vocation seriously and have proven themselves quite adept at the job. When the command is given they fire simultaneously and both inflict lethal wounds on the prisoner. Although there are difficulties associated with a proper description of states of affairs of this kind, I think it is fair to say that both men are responsible (at least in part) for its being the case that

- (Q) The prisoner is shot to death.

Now consider the fact that *Q* is identical to

- (T) The prisoner is shot to death by Kersey or the prisoner is shot to death by Nells or the prisoner is shot to death by someone else.

Kersey is not responsible for the fact that the prisoner is shot to death by Nells (or someone else for that matter). But given the fact that *Q* is identical to *T* and that Kersey is responsible (at least in part) for its being the case that *Q*, we can conclude that Kersey is responsible (at least in part) for its being the case that *T*. Therefore, assumption (i) is false. In those instances where one is responsible for its being the case that *P*, one can be responsible (at least in part) for its being the case that *P* or *R*, even though one is not responsible for its being the case that *R*. This is because the obtaining of such states of affairs may be over-determined, and responsibility for such states of affairs can be shared.

Returning to our original case, Jones is responsible for its being the case that Jones's employer's funds are embezzled by Jones acting on her own. Therefore, Jones is responsible (at least in part) for its being the case that

- (M) Either Jones's employer's funds are embezzled by Jones acting on her own or $2 + 2 = 4$,

even though she obviously is not responsible for its being the case that $2 + 2 = 4$. This is so because she shares responsibility for the obtaining of *M*.³⁸ If

³⁶ See Heinaman, p. 273.

³⁷ I am aware that some readers may find appeals to cases of over-determination *prima facie* suspect, but in this instance an appeal to such a case has seemed necessary precisely because van Inwagen's argument rests on an over-determined state of affairs, namely, *M* in my parallel exposition of his argument.

³⁸ I think we might say this is so, even though it's not at all clear who she shares responsibility with in this case. The strangeness of *M* lies in the fact that its obtaining is over-determined, not by the causally

Jones can share responsibility for the obtaining of M, then M is not an obstacle to ascribing responsibility to Jones for its being the case that L, and consequently, O (that her employer's funds are embezzled). Thus van Inwagen has not shown that our original intuitions concerning Jones were incorrect. That is, it still appears that Jones is responsible for its being the case that her employer's funds are embezzled, even though this state of affairs is inevitable.

Frankfurt-style examples attempt to show that the principle of alternate possibilities is false – that an agent lacking access to alternatives can be morally responsible. In this section I have discussed two versions of a strategy for defending the principle of alternate possibilities (or some relative of it) against the challenge posed by such examples and have concluded that these attempts fail. Furthermore, I suggest that other versions of Strategy 1 are likely to share similar faults. That is, in attempting to identify something in a Frankfurt-style example for which the agent is responsible and to which she had an alternative (step two) a proponent of Strategy 1 is likely to narrow the notion of alternative possibilities to such an extent that it will no longer be plausible to suppose that our intuitions about the agent's responsibility in a Frankfurt-style case are grounded in the agent's access to such meager alternatives. Secondly, I suggest that in the process of trying to show that the agent in such examples is not responsible for that to which she has no alternative (step three) it will be difficult for a proponent of Strategy 1 to avoid an illegitimate reliance on her conviction that responsibility does require access to alternatives. Certainly my arguments against Naylor's and van Inwagen's versions of Strategy 1 do not amount to a demonstration of these generalizations, but I suspect that the structure of the strategy itself promotes such faults.

IV

Some have suggested that incompatibilists need not rely on the principle that responsibility requires access to alternatives. That is, adopting Strategy 2, one might argue that Frankfurt-style examples are not sufficient to show that responsibility is compatible with determinism, even if they are successful in showing that access to alternatives are not required for responsibility. Perhaps determinism threatens responsibility, not because it eliminates access to alternatives, but because the way it operates in the actual sequence of events leading to an action is incompatible with responsibility.³⁹

sufficient actions of two independent parties (as in the firing squad case), but by the causally sufficient action of Jones and a necessary (and thus logically sufficient) mathematical truth. It is this strangeness of M, I believe, which prevents one from recognizing that Jones retains some responsibility for the obtaining of this over-determined state of affairs.

³⁹ My account of this strategy follows Fischer's proposal in 'Responsibility and Control', pp. 33–39. However, a similar proposal seems to be made by Robert Heinaman (though more obliquely) in

(a) Actual Sequence Incompatibilism

Note that Frankfurt-style cases neither presuppose the truth of determinism, nor that of indeterminism.⁴⁰ Suppose for example that Jones is indeterministically caused to embezzle from her employer. That is, Jones chooses on the basis of her own beliefs and desires but in such a way that the choice is not causally necessitated by them. Given this scenario, the incompatibilist can agree with Frankfurt that Jones is responsible for the embezzling even though she could not have done otherwise. On the other hand, suppose that determinism holds, and consequently Jones's choice is causally necessitated. In this case, the incompatibilist can argue that Frankfurt-style examples fail to provide any evidence that an agent could be responsible for her actions when those actions are causally determined, since in this scenario the agent performs her action *because* she could not have done otherwise, i.e. the factors which bring about the agent's action are the same factors which make it impossible that she could have done otherwise.⁴¹ Here, one might argue, it is not the absence of alternative possibilities but instead the way determinism operates in the actual sequence of events leading to an action which undermines responsibility.

To see this latter point, consider the distinction drawn earlier between alternative and actual sequence compulsion. Frankfurt-style cases involve a form of alternative sequence compulsion, but determinism, unlike Green, would operate in the actual sequence leading to Jones's decision. It is thus open to the incompatibilist to argue that Jones is not responsible for embezzling from her employer if determined, because determinism would be a form of actual sequence compulsion. Again, moral responsibility is undermined by determinism not because agents lack alternatives, but because this lack of alternatives is a sign of a deeper defect, a defect in the actual sequence leading to the agent's action.⁴² An 'actual sequence incompatibilist' can thus argue that even if the Frankfurt-style examples do show that alternatives are not necessary for free action and responsibility, this does not entail that free action and responsibility are compatible with determinism.

'Incompatibilism Without the Principle of Alternative Possibilities', pp. 275–276. Also, David Blumenfeld has proposed that incompatibilists modify the principle of alternative possibilities so as to read – One is not morally responsible for what one has done, if one did it because one could not have done otherwise. Blumenfeld's interpretation of this modified principle suggests a view similar to Fischer's. See David Blumenfeld, 'The Principle of Alternate Possibilities', *Journal of Philosophy* 67 (1971), 339–345. Finally, see also Eleonore Stump, 'Intellect, Will, and the Principle of Alternate Possibilities' in *Perspectives on Moral Responsibility*, ed. John Martin Fischer and Mark Ravizza (Ithaca: Cornell University Press, 1993), pp. 237–262, for an attempt to give an incompatibilist account of free will which does not rely on the principle of alternate possibilities.

⁴¹ Blumenfeld, pp. 341–344.

⁴² Fischer, 'Responsibility and Control', pp. 33–34.

(b) Evaluating Actual Sequence Incompatibilism

Actual sequence incompatibilism has great initial appeal. It has the benefit of recognizing the force of the Frankfurt-style cases – that what is significant for ascriptions of moral responsibility is what happens in the actual sequence as opposed to some alternative sequence – while being able to resist the contention that this shows that responsibility is compatible with determinism. The great difficulty facing the proponent of such a strategy, however, can be seen when one contrasts it with the more traditional argument for incompatibilism.

A central tenet of traditional incompatibilism is that moral responsibility requires access to alternative possibilities, and this access is to be understood as a robust ability, at the time in question, to both bring about some action or refrain from doing so.⁴³ Determinism would make access to such alternatives impossible. Thus the incompatibilist concludes: if causal determinism is true, no agent is ever morally responsible for his actions.

This is *prima facie* a powerful argument. The actual sequence incompatibilist, it appears, is unable to offer a similarly powerful argument in support of the incompatibility of determinism and responsibility. The strategy of the actual sequence incompatibilist rests on two simple claims: (1) actual sequence compulsion undermines moral responsibility; (2) determinism is a form of actual sequence compulsion. While there may be little disagreement about the first claim, the second is surely contentious. The actual sequence incompatibilist must do more than simply assert that determinism is a form of actual sequence compulsion. If the view is to be defensible, she must be able to point to some characteristic of deterministic causal sequences in virtue of which such sequences constitute a form of actual sequence compulsion. This is a considerable hurdle, for when one considers those features which seem characteristic of uncontroversial cases of actual sequence compulsion (e.g. brainwashing and hypnosis) it is not at all clear why such features should be thought to be necessary characteristics of deterministic causal sequences leading to human action.⁴⁴ To put the point

⁴³ As I pointed out in the last section, this ability traditionally has been understood in such a way that one's refraining (accomplished perhaps through the pursuit of some other course of action), like one's performance of the action in question, would follow from an intention formed on one's own.

⁴⁴ Consider two possible candidates: (1) the action results from irresistible desires (Fischer, 'Responsibility and Control', p. 34); and/or (2) the action follows from a decision or volition which is not, in the appropriate sense, the agent's own (Fischer, p. 34; Stump, p. 245). First, not only does it seem implausible to suppose that, given causal determinism, there would be no relevant difference between the desires motivating the actions of, say, a crazed heroin addict and those motivating me when I decide whether to have lunch or keep working, but also Stump has observed that there are some cases where we wish to attribute responsibility to an agent, even when the agent's action followed from irresistible desires (Stump, pp. 245–248). The notion of an agent's decision or volition being one's own is more elusive, but most recent accounts of the notion seem to be compatible with determinism. This seems to be so even of Stump's account of ownership: 'an agent's volition is his own only if his intellect represents what is willed as the good to be pursued (at that time, under some description), and the agent forms the corresponding volition

another way, why should we think that if causal determinism obtains, we would no longer be able to distinguish the behaviour of a hypnotized agent and that of an individual whose action followed from the exercise of his unimpaired practical reason?

Again, unlike those who endorse the principle of alternate possibilities, the actual sequence incompatibilist may not appeal to the fact that both actions performed as a result of actual sequence compulsion and those resulting from deterministic causal sequences are inevitable, for she has already conceded that one can be responsible for actions to which one had no alternative. Barring access to an argument such as this, it is difficult to see how one can show that causal determinism would constitute a form of actual sequence compulsion. And in the absence of an argument for this crucial claim, the initial appeal of actual sequence incompatibilism wanes considerably.

v

(a) Summary Thus Far

I have argued that the principle of alternate possibilities plays a central role in the free will defence and that in virtue of this its tenability depends on whether Frankfurt-style examples succeed in showing that acting freely and being responsible do not require access to alternatives. If this central tenet of incompatibilism is false, a major reason for denying compatibilist accounts of freedom and responsibility will have been removed. And if one grants the truth of compatibilism, the free will defence becomes vulnerable to a seemingly fatal query⁴⁵: If agents can act freely and be responsible for their actions even though they are determined to act as they do, then why did God not create a world populated with free agents who were perfectly good, i.e. who always acted rightly?

This conclusion has precipitated an extended discussion of two ways incompatibilists about moral responsibility and causal determinism have attempted to meet the challenge posed by Frankfurt-style examples. According to the first, the examples fail to show that an agent can be responsible for what she has done if she lacks freedom to do otherwise. The second sort of response to Frankfurt-style examples grants that access to alternative possibilities is not necessary for responsibility but argues that this does not entail the compatibility of responsibility and determinism. I argued that these incompatibilist strategies fail to undermine the force of Frankfurt-style

in consequence of that representation on the part of his intellect' (*Ibid.* p. 245). I see no reason why this condition could not be satisfied if causal determinism obtained. Fischer, who in his recent works has developed a compatibilist account of freedom and responsibility, mounts a similar critique of actual sequence incompatibilism in *The Metaphysics of Free Will*, pp. 147–154.

⁴⁵ Martin Davies makes an interesting, but I think unsuccessful, attempt to challenge the assumption that the free will defence is untenable if freedom is compatible with determinism. See 'Determinism and Evil', *Australasian Journal of Philosophy* 58 (1980), 116–127.

examples and thus that the examples provide persuasive support for the thesis that agents can act freely and be responsible even though determined to act as they do. Consequently, they also provide a basis for the contention that God could have eliminated the possibility of moral evil without forgoing the great good of creating creatures who act freely and are responsible for their actions. That is, contrary to the free will defender's claim, God could have created free agents who always acted rightly.

(b) *God in the Role of Green*

The main thesis of this paper has been that Frankfurt-style examples make the free will defence vulnerable to an important objection by lending support for compatibilist accounts of acting freely and being responsible. Before closing, however, I want briefly to suggest that Frankfurt-style examples may pose a threat to the free will defence, even if one believes that the incompatibilist strategies crafted in response to them are successful. By refining the incompatibilist position to meet the force of such examples, it appears that incompatibilists may have opened the free will defence to yet another version of Flew's objection.

Suppose, for example, that van Inwagen's (or Naylor's) strategy of responding to Frankfurt-style examples is successful. That strategy involves narrowing the notion of freedom to do otherwise in such a way that an agent can be morally responsible for his action, even though there is no alternative sequence in which he performs an action as the result of an intention he has formed on his own. Utilizing this notion of freedom to do otherwise, the free will defender might claim, much as before, that moral evil is the result of agents exercising their God-given freedom wrongly and that in choosing to create free agents who were responsible for their actions, it was necessary that God allow them access to such alternatives (understood now in van Inwagen or Naylor's narrower sense).

But if it is necessary that agents be able to do otherwise in only the very restricted sense van Inwagen (or Naylor) proposes, a reformulated version of the skeptic's query can yet again be raised: Why did God not create a world arranged like a series of Frankfurt-style examples in which God, occupying the role of Green, guarantees that agents always act rightly? In such a world, agents could choose, in whatever indeterministic manner the free will defender wishes, to act rightly and do so. But if an agent showed any sign of choosing to act wrongly, God would intervene to insure that he choose and act rightly.⁴⁶ Such agents could perform at least some right actions freely, in

⁴⁶ This suggestion arose in a discussion of related ideas with John Martin Fischer. It differs from the objection made by Flew in that it does not assume the truth of compatibilism. It also differs from an interesting proposal by Steven Boer in 'The Irrelevance of the Free Will Defence', *Analysis* 38 (1978), 110–112, because God's intervention takes place before the agent's choice and does not involve coincidence miracles. It is similar in some respects to some of Martin Davies's imagined scenarios (pp. 121–125) and bears some semblance to John Bishop's proposal in 'Compatibilism and the Free Will Defence' (see my note 8), which I discovered during the preparation of this manuscript. But unlike Bishop's argument,

the sense required by van Inwagen (or Naylor), and thus be responsible for those actions. Moreover, there is no apparent reason why agents should have to perform all their actions freely, in the required sense, in order for their existence to be a good thing.⁴⁷ It seems that God would have reason to create such a world, and since such a world would be free of moral evil, it is arguably preferable to one such as our own.

There is, I think, little benefit in appealing to an actual sequence version of incompatibilism in order to avoid this challenge to the free will defence. In the imagined scenario, agents can choose freely in an indeterministic manner to act rightly. They maintain control in the actual sequence of events in such instances and thus are not victims of actual sequence compulsion. They therefore appear to be responsible for what they do when they choose on their own to act rightly. The fact that God will intervene if agents show signs of choosing wrongly does not affect their responsibility for those right actions they perform freely. Nor, as I noted earlier, does the fact that agents would not be responsible for those right actions performed as a result of God's intervention eliminate the value of such agents.

There seems, however, to be an obvious objection to the scenario I have just sketched. One might object that if God intervenes upon noticing that an agent is about to choose wrongly, this would surely very quickly infect an agent's ability to freely choose rightly. For the agents in such a world would soon realize that whenever they began to form an intention to act wrongly they would be overcome by an irresistible urge to act rightly. And this realization would certainly destroy their ability to deliberate and act freely.⁴⁸ But this objection wrongly assumes that God's intervention must result in behavior analogous to that of a kleptomaniac, i.e. via the intrusion of unmotivated desires and beliefs. Alternatively, God might create agents and the world in such a way that while it is possible for them to acquire some reason to act wrongly, it will always be the case that the most reasonable thing for the agent to do, given her mental states, is that which is right.⁴⁹ In this case, if God does intervene it will not involve the intrusion of unmotivated beliefs and desires. Instead, upon reflecting on her behaviour (in those cases where God intervenes), she will view herself as having been initially inclined toward acting wrongly but as having, in the end, decided that performing the morally wrong action was not, all things considered, the best

my argument does not begin with the assumption that the Frankfurt-style examples are successful (Bishop, p. 108) but instead is motivated by the refinements incompatibilists have made to their views in light of the Frankfurt-style examples.

⁴⁷ Presumably, even a libertarian would not claim that our current situation is such that we possess freedom to do otherwise with respect to all our actions. In fact, van Inwagen has argued that agents only very rarely possess the ability to act otherwise. See 'When is the Will Free?' in *Agents, Causes, and Events*, ed. Timothy O'Connor (New York: Oxford University Press, 1995), pp. 219–238.

⁴⁸ This claim no doubt rests on the assumption that an agent must believe that she has alternatives or at least not know which of two alternatives she is considering is, in fact, the one she will choose.

⁴⁹ I owe this helpful suggestion to John Bishop, pp. 106–110, esp. fn. 2.

thing for her to do. Thus her experience would not be unlike our own when we are perhaps motivated strongly to do some wrong but succeed in acting rightly because we see that doing so is in accordance with our better judgement. If God were to intervene in such a manner, then it seems that an agent's ability to deliberate and act freely with respect to other actions could remain intact.

Let me conclude with a remark about what follows from arguments like these which question the success of the free will defence. I do not think, as Flew and Mackie seem to have thought, that belief in the existence of God hangs in the balance. Instead what my arguments suggest, if they are sound, is that theistic philosophers may need to develop more theologically robust versions of the free will defence – versions which say more about why it might have been worthwhile for God to create a world inhabited by free agents (in the libertarian sense), even though this entailed (or would most likely result in) the existence of great moral evils.⁵⁰ For if I am correct, simply claiming that it is good that there exist agents who exercise freedom with respect to morally significant actions is insufficient to the task.⁵¹

*Department of Philosophy,
University of California, Riverside,
Riverside, CA 92521-0201*

⁵⁰ As far as I know, John Hick was the first to recognize this in the first round of responses to the objections raised by Flew and Mackie. Hick in essence granted the initial force of the objections but argued that a libertarian notion of freedom was necessary in order that agents could enter freely into a genuine personal relationship with God. See *Evil and the God of Love*, rev. edn (New York: Harper and Row, 1977), pp. 271–275. John Bishop entertains a similar revision to the free will defence but seems less sure of its effectiveness, see pp. 115–119.

⁵¹ I am grateful to John M. Fischer, Paul Hoffman, Andy Coats, Avery Fouts, and Andrew Letendre for helpful comments on earlier drafts of this paper.